RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10 521,103

Source: PCT-1

Date Processed by STIC: 1-25-05

ENTERED

PCT

RAW SEQUENCE LISTING   DATE: 01/25/2005
PATENT APPLICATION: US/10/521,103   TIME: 09:03:49

Input Set : **A:\Seqlist.txt**
Output Set: **N:\CRF4\01252005\J521103.raw**

```
    4 <110> APPLICANT: Griffiths, Steven Gareth
    5         Ritchie, Rachel Jane
    6         Simard, Nathalie C.
    8 <120> TITLE OF INVENTION: Hsp70 from Arthrobacter
   11 <130> FILE REFERENCE: H-32534-A
C-->13 <140> CURRENT APPLICATION NUMBER: US/10/521,103
C-->13 <141> CURRENT FILING DATE: 2005-01-11
   13 <150> PRIOR APPLICATION NUMBER: PCT/EP03/07602
   14 <151> PRIOR FILING DATE: 2003-07-14
   16 <150> PRIOR APPLICATION NUMBER: GB 0216414.3
   17 <151> PRIOR FILING DATE: 2002-07-15
   19 <160> NUMBER OF SEQ ID NOS: 7
   21 <170> SOFTWARE: FastSEQ for Windows Version 4.0
   24 <210> SEQ ID NO: 1
   25 <211> LENGTH: 2464
   26 <212> TYPE: DNA
   27 <213> ORGANISM: Arthrobacter
   29 <400> SEQUENCE: 1
   31 ctgcgaatgt ccacgtggtg cgtgcagtga tgcgcttgaa gggcatcgca ccgcgctgaa 60
   33 ccgggttcga cccggtccca ctgagttcgc caactgagtg ggacaagccc gttctgtccc 120
   35 agtcacgcgg tcgactcagt gggaccacgc cgcagcgcga tcgatggtcg ccgcacagct 180
   37 ttttccaaag ttgagcacag gtggctcaac ttagacttga cattggtcgg ctcaagcgta 240
   39 aagttgatat cagaacactc aacttgtaag aaatcccgaa aggaaaaaac atgtcacgtg 300
   41 cagtaggcat cgacctcgga accaccaact cggtggtttc cgtcctcgaa ggcggcgagc 360
   43 ccgtcgtcat cgcgaacgcc gaaggcggcc gcaccacccc ctcagtcgtc gcgttctcca 420
   45 agagcggtga agtcctggtc ggcgaaatcg ccaagcgcca ggccgtcaac aacatcgatc 480
   47 gcaccatcgc ctcggtcaag cgccacatgg caccgactg accgtcggc atcgacgaca 540
   49 agaagtacac cgcgcaggaa atctccgccc gcaccctgat gaagctcaag aacgacgccg 600
   51 agtcctactt gggcgaaaag gtcaccgacg cggtgatcac ggttcctgcc tacttcaacg 660
   53 acgccgagcg ccaggccacc aaagaagccg gtgagatcgc cggcctgaac gtgctgcgca 720
   55 tcgtcaacga gcccactgcg gcggcgctgg cctatggctt ggacaaaggc aaagaagacg 780
   57 aactcatcct ggtcttcgac ctcggtggcg gcaccttcga cgtctcgctg ctggaagtcg 840
   59 gcaaagacga cgacggcttc tccacgatcc aggtccgcgc cacctccggc gacaaccgcc 900
   61 tgggcggcga cgactgggat cagcggatcg tcgactactt gctgaaccag ctcaaggtca 960
   63 agggcatcga cctctccaag gacaagatcg cgctgcagcg tctgcgcgaa gcttccgagc 1020
   65 aggccaagaa ggaactctcc tcggccacca gcaccaacat ctcgctgcag tacctctcgg 1080
   67 tcacccctga cggtccggtg cacttggacg agcagctgac ccgggcgaag ttccaggaac 1140
   69 tgaccgctga tctgctcgag cgcaccaaga agccgttcca ggacgtgatc gccgaggccg 1200
   71 ggatcaaggt ttccgacatc gaccacatcg tgctggtcgg cggttccacc cggatgcccg 1260
   73 cagtgaccga attggtcaag cagctggccg gtggcaagga gccgaacaag ggcgtcaacc 1320
   75 cggacgaggt ggtcgccgtc ggcgccgcgc tgcaagccgg cgtgctcaag gcgaacgca 1380
   77 aagacgtgct gctcatcgac gtcaccccgc tttccctcgg catcgaaacc aagggcggcg 1440
   79 tgatgaccaa gctgatcgag cggaacaccg cgattccgac caagcggtcc gagaccttca 1500
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/521,103

DATE: 01/25/2005
TIME: 09:03:49

Input Set : A:\Seqlist.txt
Output Set: N:\CRF4\01252005\J521103.raw

```
 81 ccacggcgga cgacaaccag ccttcggtgg ccatccaggt gttccaaggc gagcgcgagt 1560
 83 tcacccggga caacaagccg ttgggcacct tcgaactgac cggcatcgca ccggctccgc 1620
 85 gcggcgtgcc gcaggtcgaa gtcaccttcg acatcgacgc caacggcatc gtgcacgtgt 1680
 87 cggccaaaga caagggcacc ggcaaggagc agtcgatgac catcaccggc ggttcctcgc 1740
 89 tgtccaagga agacatcgag cgcatggtcg ccgacgccga ggcacacgct gcagaggaca 1800
 91 aggcccggcg cgagcaggcc gaggcccgca acagcgccga gcagctggcg tactcggtgg 1860
 93 acaagatcct caccgacaat gacgacaagc tgccggaaga ggtcaagacg gaggtcaagg 1920
 95 ccgacgtcgg ggcgctcaag accgcgctgg ccggcaccga tgaggacgcg gtcgaggcgg 1980
 97 cctcggagaa gctgcaggct tcgcagacca aactcggcgg agcgatttac gcttcggccc 2040
 99 aggccgaggg tgccgccgct gccggtgacg ccccgagcga aggtgccaag cccgacgaag 2100
101 acatcgtcga cgccgagatc gtggacgaag aagaaccgaa gaacgagaag aagtagtcat 2160
103 gtccgaccag agccaatctg atcagggccg caaccccgaa aaagacgaaa ccgacgtgga 2220
105 cccggcaacg ggtcccgccg gggacgttcc ggaggagcag gatcctttgg cgcaagtcga 2280
107 agacatcctg aacaatgccg aggtgcccgc cgacgagtcg gtggcccagg gcgccgggca 2340
109 ggtggatgcc gcagaactca gaacgatct gctgcgcttg caggccgaat acgtgaacta 2400
111 ccgcaaacgc gtcgagcggg acaccagccc gggccgtcga ccacgcgtgc cctatagtaa 2460
113 gggc                                                           2464
117 <210> SEQ ID NO: 2
118 <211> LENGTH: 621
119 <212> TYPE: PRT
120 <213> ORGANISM: Arthrobacter
122 <400> SEQUENCE: 2
124 Met Ser Arg Ala Val Gly Ile Asp Leu Gly Thr Thr Asn Ser Val Val
125  1               5                  10                  15
127 Ser Val Leu Glu Gly Gly Glu Pro Val Val Ile Ala Asn Ala Glu Gly
128             20                  25                  30
130 Gly Arg Thr Thr Pro Ser Val Val Ala Phe Ser Lys Ser Gly Glu Val
131             35                  40                  45
133 Leu Val Gly Glu Ile Ala Lys Arg Gln Ala Val Asn Asn Ile Asp Arg
134         50                  55                  60
136 Thr Ile Ala Ser Val Lys Arg His Met Gly Thr Asp Trp Thr Val Gly
137 65                  70                  75                  80
139 Ile Asp Asp Lys Lys Tyr Thr Ala Gln Glu Ile Ser Ala Arg Thr Leu
140                 85                  90                  95
142 Met Lys Leu Lys Asn Asp Ala Glu Ser Tyr Leu Gly Glu Lys Val Thr
143             100                 105                 110
145 Asp Ala Val Ile Thr Val Pro Ala Tyr Phe Asn Asp Ala Glu Arg Gln
146             115                 120                 125
148 Ala Thr Lys Glu Ala Gly Glu Ile Ala Gly Leu Asn Val Leu Arg Ile
149         130                 135                 140
151 Val Asn Glu Pro Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Gly
152 145                 150                 155                 160
154 Lys Glu Asp Glu Leu Ile Leu Val Phe Asp Leu Gly Gly Gly Thr Phe
155                 165                 170                 175
157 Asp Val Ser Leu Leu Glu Val Gly Lys Asp Asp Asp Gly Phe Ser Thr
158             180                 185                 190
160 Ile Gln Val Arg Ala Thr Ser Gly Asp Asn Arg Leu Gly Gly Asp Asp
161         195                 200                 205
163 Trp Asp Gln Arg Ile Val Asp Tyr Leu Leu Asn Gln Leu Lys Val Lys
```

RAW SEQUENCE LISTING   DATE: 01/25/2005
PATENT APPLICATION: US/10/521,103   TIME: 09:03:49

Input Set : A:\Seqlist.txt
Output Set: N:\CRF4\01252005\J521103.raw

```
164     210                    215                    220
166 Gly Ile Asp Leu Ser Lys Asp Lys Ile Ala Leu Gln Arg Leu Arg Glu
167 225                    230                    235                    240
169 Ala Ser Glu Gln Ala Lys Lys Glu Leu Ser Ser Ala Thr Ser Thr Asn
170                    245                    250                    255
172 Ile Ser Leu Gln Tyr Leu Ser Val Thr Pro Asp Gly Pro Val His Leu
173                    260                    265                    270
175 Asp Glu Gln Leu Thr Arg Ala Lys Phe Gln Glu Leu Thr Ala Asp Leu
176             275                    280                    285
178 Leu Glu Arg Thr Lys Lys Pro Phe Gln Asp Val Ile Ala Glu Ala Gly
179             290                    295                    300
181 Ile Lys Val Ser Asp Ile Asp His Ile Val Leu Val Gly Gly Ser Thr
182 305                    310                    315                    320
184 Arg Met Pro Ala Val Thr Glu Leu Val Lys Gln Leu Ala Gly Gly Lys
185                    325                    330                    335
187 Glu Pro Asn Lys Gly Val Asn Pro Asp Glu Val Val Ala Val Gly Ala
188             340                    345                    350
190 Ala Leu Gln Ala Gly Val Leu Lys Gly Glu Arg Lys Asp Val Leu Leu
191             355                    360                    365
193 Ile Asp Val Thr Pro Leu Ser Leu Gly Ile Glu Thr Lys Gly Gly Val
194       370                    375                    380
196 Met Thr Lys Leu Ile Glu Arg Asn Thr Ala Ile Pro Thr Lys Arg Ser
197 385                    390                    395  ·                 400
199 Glu Thr Phe Thr Thr Ala Asp Asp Asn Gln Pro Ser Val.Ala Ile Gln
200                    405                    410                    415
202 Val Phe Gln Gly Glu Arg Glu Phe Thr Arg Asp Asn Lys Pro Leu Gly
203             420                    425                    430
205 Thr Phe Glu Leu Thr Gly Ile Ala Pro Ala Pro Arg Gly Val Pro Gln
206             435                    440                    445
208 Val Glu Val Thr Phe Asp Ile Asp Ala Asn Gly Ile Val His Val Ser
209       450                    455                    460
211 Ala Lys Asp Lys Gly Thr Gly Lys Glu Gln Ser Met Thr Ile Thr Gly
212 465                    470                    475                    480
214 Gly Ser Ser Leu Ser Lys Glu Asp Ile Glu Arg Met Val Ala Asp Ala
215                    485                    490                    495
217 Glu Ala His Ala Ala Glu Asp Lys Ala Arg Arg Glu Gln Ala Glu Ala
218             500                    505                    510
220 Arg Asn Ser Ala Glu Gln Leu Ala Tyr Ser Val Asp Lys Ile Leu Thr
221             515                    520                    525
223 Asp Asn Asp Asp Lys Leu Pro Glu Glu Val Lys Thr Glu Val Lys Ala
224       530                    535                    540
226 Asp Val Gly Ala Leu Lys Thr Ala Leu Ala Gly Thr Asp Glu Asp Ala
227 545                    550                    555                    560
229 Val Glu Ala Ala Ser Glu Lys Leu Gln Ala Ser Gln Thr Lys Leu Gly
230             565                    570                    575
232 Gly Ala Ile Tyr Ala Ser Ala Gln Ala Glu Gly Ala Ala Ala Ala Gly
233             580                    585                    590
235 Asp Ala Pro Ser Glu Gly Ala Lys Pro Asp Glu Asp Ile Val Asp Ala
236             595                    600                    605
```

RAW SEQUENCE LISTING   DATE: 01/25/2005
PATENT APPLICATION: US/10/521,103   TIME: 09:03:49

Input Set : **A:\Seqlist.txt**
Output Set: **N:\CRF4\01252005\J521103.raw**

```
    238 Glu Ile Val Asp Glu Glu Glu Pro Lys Asn Glu Lys Lys
    239     610             615             620
    243 <210> SEQ ID NO: 3
    244 <211> LENGTH: 20
    245 <212> TYPE: PRT
    246 <213> ORGANISM: Arthrobacter
    248 <400> SEQUENCE: 3
    250 Met Ser Arg Ala Val Gly Ile Asp Leu Gly Thr Thr Asn Ser Val Val
    251 1               5               10              15
    253 Ser Val Leu Glu
    254             20
    258 <210> SEQ ID NO: 4
    259 <211> LENGTH: 20
    260 <212> TYPE: DNA
    261 <213> ORGANISM: Artificial Sequence
    263 <220> FEATURE:
    264 <223> OTHER INFORMATION: degenerate primers for cloning Arthrobacter hsp70
W--> 266 <221> NAME/KEY: misc_feature
    267 <222> LOCATION: 6, 18
    268 <223> OTHER INFORMATION: n = A,T,C or G
W--> 270 <221> misc_feature
    271 <222> LOCATION: 15
    272 <223> OTHER INFORMATION: v= A, G, or C
W--> 274 <400> 4
W--> 276 gtcggnatcg acctvggnac                                    20
    280 <210> SEQ ID NO: 5
    281 <211> LENGTH: 17
    282 <212> TYPE: DNA
    283 <213> ORGANISM: Artificial Sequence
    285 <220> FEATURE:
    286 <223> OTHER INFORMATION: degenerate primers for cloning Arthrobacter hsp70
W--> 288 <221> NAME/KEY: misc_feature
    289 <222> LOCATION: 6
    290 <223> OTHER INFORMATION: s = C or G
W--> 292 <400> 5
    294 gcggtsggct cgttgac                                        17
    298 <210> SEQ ID NO: 6
    299 <211> LENGTH: 20
    300 <212> TYPE: DNA
    301 <213> ORGANISM: Artificial Sequence
    303 <220> FEATURE:
    304 <223> OTHER INFORMATION: degenerate primers for cloning Arthrobacter hsp70
W--> 306 <221> NAME/KEY: misc_feature
    307 <222> LOCATION: 6, 9, 18
    308 <223> OTHER INFORMATION: n = A,T,C or G
W--> 310 <221> misc_feature
    311 <222> LOCATION: 3, 12
    312 <223> OTHER INFORMATION: r = A or G
W--> 314 <221> misc_feature
```

```
      315 <222> LOCATION: 15
      316 <223> OTHER INFORMATION: y = C or T
W--> 318 <400> 6
W--> 320 cargcncana argaygcngg                                    20
      324 <210> SEQ ID NO: 7
      325 <211> LENGTH: 20
      326 <212> TYPE: DNA
      327 <213> ORGANISM: Artificial Sequence
      329 <220> FEATURE:
      330 <223> OTHER INFORMATION: degenerate primers for cloning Arthrobacter hsp70
W--> 332 <221> NAME/KEY: misc_feature
      333 <222> LOCATION: 3, 6, 15
      334 <223> OTHER INFORMATION: n = A,T,C or G
W--> 336 <221> misc_feature
      337 <222> LOCATION: 9
      338 <223> OTHER INFORMATION: y = C or T
W--> 340 <221> misc_feature
      341 <222> LOCATION: 12, 18
      342 <223> OTHER INFORMATION: r = A or G
W--> 344 <400> 7
W--> 346 gcncangcyt crtcnggrtt                                     20
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/25/2005
PATENT APPLICATION: **US/10/521,103**   TIME: 09:03:50

Input Set : **A:\Seqlist.txt**
Output Set: **N:\CRF4\01252005\J521103.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; N Pos. 6,18
Seq#:6; N Pos. 6,9,18
Seq#:7; N Pos. 3,6,15

VERIFICATION SUMMARY   DATE: 01/25/2005
PATENT APPLICATION: **US/10/521,103**   TIME: 09:03:50

Input Set : **A:\Seqlist.txt**
Output Set: **N:\CRF4\01252005\J521103.raw**

```
L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:266 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:270 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:4
L:274 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:4
L:276 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:288 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:292 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:5
L:306 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:310 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:6
L:314 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:6
L:318 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:6
L:320 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:332 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:336 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:7
L:340 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:7
L:344 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:7
L:346 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
```